A SEQUENTIAL BAYESIAN INFERENCE FRAMEWORK FOR BLIND FREQUENCY OFFSET ESTIMATION

Theodoros Tsiligkaridis, Keith W. Forsythe

MIT Lincoln Laboratory
Lexington, MA 02421 USA
{ttsili,forsythe}@ll.mit.edu

ABSTRACT

Precise estimation of synchronization parameters is essential for reliable data detection in digital communications and phase errors can result in significant performance degradation. The literature on estimation of synchronization parameters, including the carrier frequency offset, are based on approximations or heuristics because the optimal estimation problem is analytically intractable for most cases of interest. We develop an online Bayesian inference procedure for blind estimation of the frequency offset, for arbitrary signal constellations. Our unified approach is built on a sequential inference procedure that leverages a novel result on conjugacy of the von Mises and Gaussian distributions. This conjugacy allows for an easily computable, closed form parametric expression for the posterior distribution of the parameters given the streaming data, in which hyperparameters are recursively updated, making the optimal sequential estimation problem mathematically tractable. Our algorithm is computationally efficient and can be implemented in real-time with very low memory requirements. Numerical experiments are also provided and show that our methods outperform heuristic sequential carrier frequency offset estimators.

Index Terms— Sequential Estimation, Blind Synchronization, Frequency Offset Estimation, Bayesian Inference, Phase tracking.

1. INTRODUCTION

Synchronization plays a critical role in attaining reliable digital transmission through wireless channels. Timing and frequency offsets need to be estimated well in order to align the received signal appropriately at the receiver before data detection. Optimal estimators for synchronization parameters do not exist and as a result, most of the existing literature uses approximate maximum-likelihood techniques and heuristics [1, 2]. Even in the simpler setting, where the amplitude is known and constant, optimal estimation of the frequency offset is not known in closed-form and only approximate maximum-likelihood (ML) and maximum a posterior (MAP) approaches are tractable [3].

In digital communications, there is often a mismatch between the local oscillator of the transmitter and the receiver. This translates to a carrier frequency error when downconverting at the receiver, which effectively rotates the signal constellation from sample-to-sample. For small carrier frequency errors, Δf, the constellation rotates slowly and tracking its rotation rate is required for successful data detection. Popular methods for carrier synchronization include the use of a phase-lock loop (PLL), which often requires hardware implementation or waveform-level block-based processing in software, both of which can be expensive. Furthermore, the PLL approach often requires pilot data for higher order modulations. A synchronization approach based on model-based sequential Monte Carlo (SMC) techniques was proposed in [4] to estimate the timing offset and the data, but is heavily dependent on known state-space dynamic models governing the evolution of the parameters, several approximations are made to make the SMC algorithm tractable, and is computationally expensive as the complexity of the SMC grows exponentially fast [4].

In contrast, our approach is not data-aided (i.e., blind), works for arbitrary signal modulations, operates at the sample-level, is simple to implement, and has very low latency and storage requirement as it proceeds on-the-fly in an online fashion. In this paper, we propose a novel sequential Bayesian inference framework to estimate the unknown carrier frequency offset and the parameters of the signal modulation.

We leverage ideas from the sequential inference approach of [5] to derive a sequential Bayesian algorithm to estimate the frequency offset and the constellation parameters given streaming non-i.i.d. received samples. This is a challenging machine learning problem primarily because of the temporal dynamics of the data generation process. Our method exploits the temporal relation from sample-to-sample and uses the online clustering and parameter estimation framework developed in [5]. Solid empirical performance is observed for slow enough rotation rates in our experiments.

2. SEQUENTIAL BAYESIAN INFERENCE FRAMEWORK

Here, we review the adaptive sequential updating and greedy search (ASUGS) framework of [5] for online clustering and parameter estimation. Define the unknown parameters $\theta = (\mu_1, \dots, \mu_K, \delta)$, where $\delta \in [0, 2\pi)$ is the unknown rotation offset and μ_h is the mean of each class (i.e., cluster center). Let the observations be given by $\mathbf{y}_i \in \mathbb{R}^d$, and γ_i to denote the class label of the ith observation (a latent variable). We define the available information at time i as $\mathbf{y}^{(i)} = \{\mathbf{y}_1, \dots, \mathbf{y}_i\}$ and $\gamma^{(i-1)} = \{\gamma_1, \dots, \gamma_{i-1}\}$. The ASUGS algorithm is as follows. Set $\gamma_1 = 1$ and $k_1 = 1$. Calculate $\pi(\theta_1|\mathbf{y}_1, \gamma_1)$. For $i \geq 2$,

1. Choose best class label for $\mathbf{y}_i$:

$$\gamma_i \sim \{q_h^{(i)}\} = \left\{ \frac{L_{i,h}(\mathbf{y}_i)\pi_{i,h}}{\sum_{h'} L_{i,h'}(\mathbf{y}_i)\pi_{i,h'}} \right\}.$$

Distribution A: Public Release. This work is sponsored by the Assistant Secretary of Defense for Re-search & Engineering under Air Force Contract #FA8721-05-C-0002. Opin-ions, interpretations, conclusions and recommendations are those of the au-thor and are not necessarily endorsed by the United States Government.

2. Update the posterior distribution using $\mathbf{y}_i, \gamma_i$:

$$\pi(\theta_{\gamma_i}|\mathbf{y}^{(i)}, \gamma^{(i)}) \propto f(\mathbf{y}_i|\theta_{\gamma_i})\pi(\theta_{\gamma_i}|\mathbf{y}^{(i-1)}, \gamma^{(i-1)}). \quad (1)$$

where θ_h are the parameters of class h, $f(\mathbf{y}_i|\theta_h)$ is the observation density conditioned on class h. The conditional likelihood $P(\mathbf{y}_i|\gamma_i = h, \mathbf{y}^{(i-1)}, \gamma^{(i-1)})$ is denoted by $L_{i,h}(\mathbf{y}_i)$ and $\pi_{i,h}$ denotes the class priors. The algorithm sequentially allocates observations $\mathbf{y}_i$ to classes by sampling the conditional posterior probability distribution $\{q_h^{(i)}\}$.

3. PROBLEM FORMULATION AND PROBABILISTIC MODEL

Consider a digital communication system where symbols from a discrete unknown alphabet, $x_m \in \mathcal{A}$ are transmitted through a channel. The received signal at the front-end of the receiver consisting of a matched filter can be modeled as:

$$y(t) = e^{j2\pi\Delta f t} \sum_m x_m g(t - mT) + w(t)$$

where T is the symbol period, $g(\cdot)$ is the raised-cosine pulse waveform, and $w(t)$ is additive white Gaussian noise (AWGN) with power spectral density $N_0/2$. Here, Δf is the carrier frequency error. Sampling the output of the matched filter at a rate $1/T$, we obtain the discrete-time signal:

$$y_k = e^{j2\pi\Delta f kT} x_k + w_k \quad (2)$$

where $y_k = y(kT), w_k = w(kT)$, and $k = 0, 1, \dots$ is the discrete-time index. Note the phase rotation offset $\delta = 2\pi\Delta f T$ is proportional to the carrier frequency error Δf.

We consider a Bayesian framework for estimating the unknown mean for each class and unknown phase rotation offset δ. This framework can be extended to include unknown covariances as well (corresponding to unknown SNR), but is not considered here for simplicity (see [5]). The means can be considered as known by setting the uncertainty to zero (see Section 4).

The observation model (2) in vector form is given by:

$$\begin{aligned} \mathbf{y}_i &= \mathbf{R}(\theta_i)\mathbf{x}_i + \mathbf{w}_i \\ \theta_{i+1} &= \theta_i + \delta \end{aligned}$$

where $\mathbf{x}_i \in \mathcal{A}$ are symbols from a constellation, $\mathbf{w}_i$ is additive Gaussian noise with covariance $\sigma^2\mathbf{I}$. Here, $\mathbf{R}(\theta)$ is a rotation matrix given by

$$\mathbf{R}(\theta) = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

Without loss of generality, we assume $\theta_1 = 0$. The unknown parameters in the model are the constellation symbols $\mathbf{x} \in \mathcal{A}$, and the phase offset δ. Let $K = |\mathcal{A}|$ denote the size of the symbol alphabet.

The probabilistic model for the unknown parameters is given as:

$$\begin{aligned} \mathbf{y}|\mu, \delta &\sim \mathcal{N}(\cdot|\mathbf{R}(\delta)\mu, \sigma^2\mathbf{I}) \\ \mu &\sim \mathcal{N}(\cdot|\mu_0, \sigma_0^2\mathbf{I}) \\ \delta &\sim \mathcal{V}(\cdot|\delta_0, \kappa_0) \end{aligned} \quad (3)$$

where $\mathcal{N}(\cdot|\mu, \mathbf{\Sigma})$ denote the multivariate normal distribution with mean μ and covariance matrix $\mathbf{\Sigma}$, and $\mathcal{V}(\cdot|\delta_0, \kappa_0)$ denotes the von Mises distribution with direction parameter δ_0 and concentration parameter κ_0. The parameters $\theta = (\mu, \delta) \in \mathbb{R}^d \times [-\pi, \pi)$ follow a normal-von Mises joint distribution. The corresponding graphical model is shown in Fig. 1.

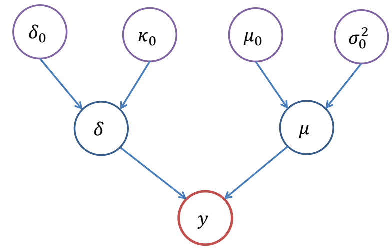

Fig. 1. Graphical model for simultaneous phase tracking and parameter estimation problem.

For concreteness, let us write the distribution in (3):

$$\begin{aligned} f(\mathbf{y}_i|\theta) = p(\mathbf{y}_i|\mu, \delta) &= \frac{1}{(2\pi\sigma^2)^{d/2}} e^{-\frac{1}{2\sigma^2}\|\mathbf{y}_i - \mathbf{R}(\delta)\mu\|_2^2} \\ p(\mu) &= \frac{1}{(2\pi\sigma_0^2)^{d/2}} e^{-\frac{1}{2\sigma_0^2}\|\mu - \mu_0\|_2^2} \\ p(\delta) &= \frac{1}{2\pi I_0(\kappa_0)} e^{\kappa_0 \cos(\delta - \delta_0)} \end{aligned}$$

where $I_0(\cdot)$ is the modified Bessel function of order 0.

Due to the conjugacy of the distributions, the posterior distribution $\pi(\theta_h|\mathbf{y}^{(i-1)}, \gamma^{(i-1)})$ always has the form:

$$\begin{aligned} \pi(\theta_h|\mathbf{y}^{(i-1)}, \gamma^{(i-1)}) &= \mathcal{N}(\mu_h|\mu_h^{(i-1)}, (\sigma_h^{(i-1)})^2\mathbf{I}) \\ &\quad \times \mathcal{V}(\delta|\delta^{(i-1)}, \kappa^{(i-1)}) \end{aligned} \quad (4)$$

where $\mu_h^{(i-1)}, \sigma_h^{(i-1)}, \delta^{(i-1)}, \kappa^{(i-1)}$ are hyperparameters that can be recursively computed as new samples come in. This factorization is proven analytically in Section 4.

We will show in Section 4 that the model (3) leads to closed-form expressions for hyperparameter updates due to conjugacy.

4. CONJUGACY AND HYPERPARAMETER UPDATES

In this section, we derive the hyperparameter updates. For simplicity, let $\mathbf{y}$ be a generic observation of the form:

$$\mathbf{y} = \mathbf{R}(\delta)\mathbf{x} + \mathbf{w} \quad (5)$$

The conditional distribution $p(\mu|\delta, \mathbf{y}) \propto p(\mathbf{y}|\mu, \delta)p(\mu)$ is multivariate normal with mean and covariance given by:

$$\begin{aligned} \mathbb{E}[\mu|\delta, \mathbf{y}] &= \frac{\sigma_0^2}{\sigma^2 + \sigma_0^2}\mathbf{R}(\delta)^T\mathbf{y} + \frac{\sigma^2}{\sigma^2 + \sigma_0^2}\mu_0 \\ \mathrm{Cov}(\mu|\delta, \mathbf{y}) &= \frac{\sigma^2\sigma_0^2}{\sigma^2 + \sigma_0^2}\mathbf{I} \end{aligned}$$

Recall the von Mises prior distribution on δ, i.e., $p(\delta) \propto e^{\kappa_0 \cos(\delta - \delta_0)}$. The posterior is given by

$$\begin{aligned} p(\delta|\mathbf{y}) &\propto p(\mathbf{y}|\delta)p(\delta) \\ &\propto \exp\left(\frac{-\|\mathbf{y} - \mathbf{R}(\delta)\mu_0\|_2^2}{2(\sigma^2 + \sigma_0^2)} + \kappa_0 \cos(\delta - \delta_0)\right) \\ &\propto \exp\left(\frac{\mathbf{y}^T\mathbf{R}(\delta)\mu_0}{\sigma^2 + \sigma_0^2} + \kappa_0 \cos(\delta - \delta_0)\right) \end{aligned} \quad (6)$$

Next, write $\mathbf{y} = [y_R, y_I]^T, \mu_0 = [\mu_{0,R}, \mu_{0,I}]^T$. Then, direct calculations yield:

$$\mathbf{y}^T \mathbf{R}(\delta)\mu_0 = (\mathbf{y}^T \mu_0)\cos\delta + (\mathbf{y}^T \tilde{\mu}_0)\sin\delta \tag{7}$$

where $\tilde{\mu}_0 = [-\mu_{0,I}, \mu_{0,R}]^T$. Furthermore, from a trigonometric identity, we obtain:

$$\kappa_0 \cos(\delta - \delta_0) = \kappa_0 \cos(\delta_0)\cos\delta + \kappa_0 \sin(\delta_0)\sin\delta \tag{8}$$

Using (7) and (8) into (6), we obtain:

$$p(\delta|\mathbf{y}) \propto \exp\Bigg(\left(\kappa_0 \cos(\delta_0) + \frac{\mathbf{y}^T\mu_0}{\sigma^2 + \sigma_0^2}\right)\cos\delta + \left(\kappa_0 \sin(\delta_0) + \frac{\mathbf{y}^T\tilde{\mu}_0}{\sigma^2 + \sigma_0^2}\right)\sin\delta \Bigg)$$

This equals $\exp(\kappa_{new}\cos(\delta - \delta_{new}))$ for all $\delta \in [-\pi, \pi)$ iff:

$$\kappa_0 \cos(\delta_0) + \frac{\mathbf{y}^T\mu_0}{\sigma^2 + \sigma_0^2} = \kappa_{new}\cos(\delta_{new})$$
$$\kappa_0 \sin(\delta_0) + \frac{\mathbf{y}^T\tilde{\mu}_0}{\sigma^2 + \sigma_0^2} = \kappa_{new}\sin(\delta_{new})$$

Solving for κ_{new} and δ_{new}, we obtain:

$$\kappa_{new}^2 = \left(\kappa_0 \cos(\delta_0) + \frac{\mathbf{y}^T\mu_0}{\sigma^2 + \sigma_0^2}\right)^2 + \left(\kappa_0 \sin(\delta_0) + \frac{\mathbf{y}^T\tilde{\mu}_0}{\sigma^2 + \sigma_0^2}\right)^2$$
$$\delta_{new} = \tan^{-1}\left(\frac{\kappa_0 \sin(\delta_0) + \frac{\mathbf{y}^T\tilde{\mu}_0}{\sigma^2+\sigma_0^2}}{\kappa_0 \cos(\delta_0) + \frac{\mathbf{y}^T\mu_0}{\sigma^2+\sigma_0^2}}\right)$$

Returning to the model at the ith time instant, pre-multiplying with the rotation matrix $\mathbf{R}(\theta_{i-1})^T$:

$$\begin{aligned}\mathbf{R}(\theta_{i-1})^T\mathbf{y}_i &= \mathbf{R}(\theta_{i-1})^T\mathbf{R}(\theta_i)\mathbf{x}_i + \mathbf{R}(\theta_{i-1})^T\mathbf{w}_i \\ &= \mathbf{R}(\theta_{i-1})^{-1}\mathbf{R}(\theta_{i-1})\mathbf{R}(\delta)\mathbf{x}_i + \mathbf{R}(\theta_{i-1})^T\mathbf{w}_i \\ &= \mathbf{R}(\delta)\mathbf{x}_i + \mathbf{R}(\theta_{i-1})^T\mathbf{w}_i\end{aligned}$$

where $\mathbf{R}(\theta_{i-1})^T\mathbf{w}_i$ is Gaussian noise with zero mean and covariance $\sigma^2\mathbf{I}$, due to rotation invariance. Thus, the model (5) becomes applicable when applied to $\mathbf{R}(\theta_{i-1})^T\mathbf{y}_i$. This essentially amounts to rotating the vector $\mathbf{y}_i$ by an angle θ_{i-1}. A graphical illustration of this rotation is shown in Fig. 2.

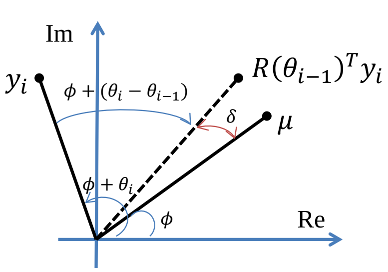

Fig. 2. Illustration of pre-rotation of observation $\mathbf{y}_i$ to reduce to model (5).

To summarize, once the γ_ith component is chosen, the parameter updates for the γ_ith class become:

$$\delta^{(i)} = \tan^{-1}\left(\frac{\kappa^{(i-1)}\sin\delta^{(i-1)} + \frac{\left\langle \mathbf{R}(\theta^{(i-1)})^T\mathbf{y}_i, \tilde{\mu}_{\gamma_i}^{(i-1)}\right\rangle}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}}{\kappa^{(i-1)}\cos\delta^{(i-1)} + \frac{\left\langle \mathbf{R}(\theta^{(i-1)})^T\mathbf{y}_i, \mu_{\gamma_i}^{(i-1)}\right\rangle}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}}\right) \tag{9}$$

$$\begin{aligned}(\kappa^{(i)})^2 = &\left(\kappa^{(i-1)}\sin\delta^{(i-1)} + \frac{\left\langle \mathbf{R}(\theta^{(i-1)})^T\mathbf{y}_i, \tilde{\mu}_{\gamma_i}^{(i-1)}\right\rangle}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}\right)^2 \\ &+ \left(\kappa^{(i-1)}\cos\delta^{(i-1)} + \frac{\left\langle \mathbf{R}(\theta^{(i-1)})^T\mathbf{y}_i, \mu_{\gamma_i}^{(i-1)}\right\rangle}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}\right)^2\end{aligned} \tag{10}$$

$$\theta^{(i)} = \theta^{(i-1)} + \delta^{(i)} \tag{11}$$

$$\gamma^{(i)} \sim q_h^{(i)} = \frac{L_{i,h}(\mathbf{R}(\theta^{(i)})^T\mathbf{y}_i)\pi_{i,h}}{\sum_h L_{i,h}(\mathbf{R}(\theta^{(i)})^T\mathbf{y}_i)\pi_{i,h}}$$
$$\mu_{\gamma_i}^{(i)} = \frac{(\sigma_{\gamma_i}^{(i-1)})^2}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}\mathbf{R}(\theta^{(i)})^T\mathbf{y}_i + \frac{\sigma^2}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}\mu_{\gamma_i}^{(i-1)} \tag{12}$$

$$(\sigma_{\gamma_i}^{(i)})^2 = (\sigma_{\gamma_i}^{(i-1)})^2\left(\frac{\sigma^2}{\sigma^2 + (\sigma_{\gamma_i}^{(i-1)})^2}\right) \tag{13}$$

For exactly known means (i.e., constellation parameters), one can set $\sigma_h^{(i)} = 0$ and $\mu_h^{(i)} = \mu_h$ for all h, i. In this case, the updates (12)-(13) become superfluous.

4.1. Calculation of Conditional Likelihood

To calculate the class posteriors $\{q_h^{(i)}\}$, the conditional likelihoods of $\bar{\mathbf{y}}_i \stackrel{\text{def}}{=} \mathbf{R}(\theta^{(i)})^T\mathbf{y}_i$ given assignment to class h and the previous class assignments need to be calculated first. The conditional likelihood of $\mathbf{y}_i$ given assignment to class h and the history $(\mathbf{y}^{(i-1)}, \gamma^{(i-1)})$ is given by:

$$L_{i,h}(\bar{\mathbf{y}}_i) = \int f(\bar{\mathbf{y}}_i|\theta_h)\pi(\theta_h|\mathbf{y}^{(i-1)}, \gamma^{(i-1)})d\theta_h \tag{14}$$

We recall from (4) that the posterior distribution has the product form:

$$\pi(\theta_h|\mathbf{y}^{(i-1)}, \gamma^{(i-1)}) = \mathcal{N}(\mu_h|\mu_h^{(i-1)}, (\sigma_h^{(i-1)})^2\mathbf{I})\mathcal{V}(\delta|\delta^{(i-1)}, \kappa^{(i-1)})$$

For large $\kappa^{(i-1)}$, which is the case after a few iterations, this can be approximated as:

$$\pi(\theta_h|\mathbf{y}^{(i-1)}, \gamma^{(i-1)}) \approx \mathcal{N}(\mu_h|\mu_h^{(i-1)}, (\sigma_h^{(i-1)})^2\mathbf{I})\mathcal{D}(\delta - \delta^{(i-1)}) \tag{15}$$

where $\mathcal{D}(\cdot)$ denotes the Dirac function. Using this approximation (15) into (14):

$$\begin{aligned}L_{i,h}(\mathbf{y}_i) &\approx \int_{\mathbb{R}^d} \mathcal{N}(\mathbf{y}_i|\mathbf{R}(\theta^{(i)})\mu, \sigma^2\mathbf{I}) \times \mathcal{N}(\mu|\mu_h^{(i-1)}, (\sigma_h^{(i-1)})^2\mathbf{I})d\mu \\ &= \frac{1}{(2\pi(\sigma^2 + (\sigma_h^{(i-1)})^2))^{d/2}} e^{-\frac{\|\mathbf{R}(\theta^{(i)})^T\mathbf{y}_i - \mu_h^{(i-1)}\|_2^2}{2(\sigma^2 + (\sigma_h^{(i-1)})^2)}}\end{aligned} \tag{16}$$

which corresponds to the Gaussian distribution $\mathbf{y}_i \sim \mathcal{N}(\cdot|\mathbf{R}(\theta^{(i)})\mu_h^{(i-1)}, (\sigma_h^{(i-1)})^2)\mathbf{I})$.

5. SIMULATIONS

In this section, we perform a simulation experiment on detecting and estimating the constellation parameters of a 16-QAM modulation. The phase rotation angle per-sample is $\delta = 3°$. The correct number of classes for this data set is 16, each one corresponding to 4 information bits.

Data symbols $\{\mathbf{x}_i\}$ are randomly (uniformly) chosen from the 16-QAM constellation and corrupted with additive noise at SNR of 15dB. The data is plotted in Fig. 3, along with the original signaling locations arranged in a rectangular grid.

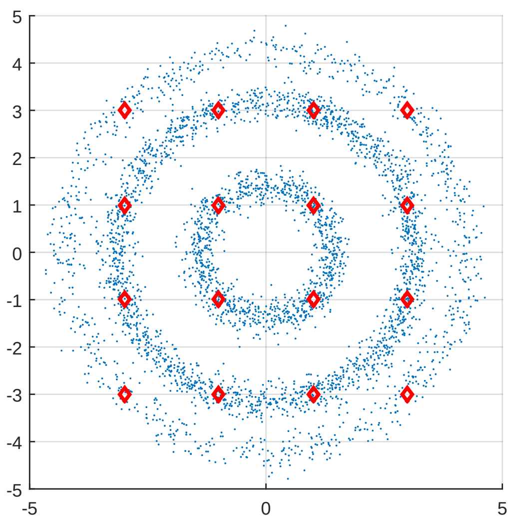

Fig. 3. Noisy constellation data from 16-QAM constellation with phase rotation rate of $\delta = 3°$ at 15dB.

For comparison, we consider the naive sequential phase estimator:

$$\theta^{(i)} = \angle\left\{\frac{y_i}{P_{\mathcal{A}}(e^{-j\theta^{(i-1)}}y_i)}\right\} \tag{17}$$

where $y_i = y_i^R + jy_i^I$, $P_{\mathcal{A}}(\mathbf{z}) := \mu_{h^*}$, $h^* = \arg\min_h \|\mathbf{z} - \mu_h\|_2^2$ is the minimum norm projection on the constellation $\mathcal{A}$. Fig. 4 shows the clustering performance of this naive method, and the clusters are very noisy and not estimated well. As shown in Fig. 4, the phase offset δ is not learned as the number of samples grow; although the mean is around $3°$, the variance is quite high and is not decaying as a function of iteration.

To alleviate this problem, we aim to use the online Bayesian estimation algorithm developed in this paper to estimate the frequency offset. Fig. 5 shows the clustering performance and the estimation performance of the Bayesian algorithm. In our simulations, we implemented the paremeter updates (9)-(11), and set $\mu_h^{(i)} = \mu_h, \sigma_h^{(i)} = 0$ for all h, i because no uncertainty in the constellation parameters was assumed. The algorithm was initialized with $\delta^{(1)} = 0$. With the Bayesian learning algorithm based on the von Mises prior, asymptotic learning occurs for the phase offset δ as more samples are processed. This leads to a significantly more stable performance when compared to the naive phase estimator (17).

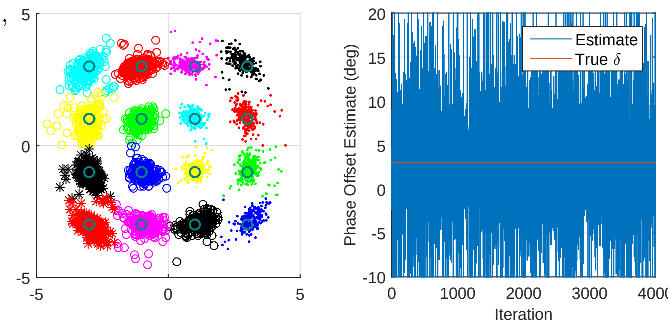

Fig. 4. Resulting compensated constellation (left) and frequency offset estimation performance (right) for naive sequential phase estimator (17).

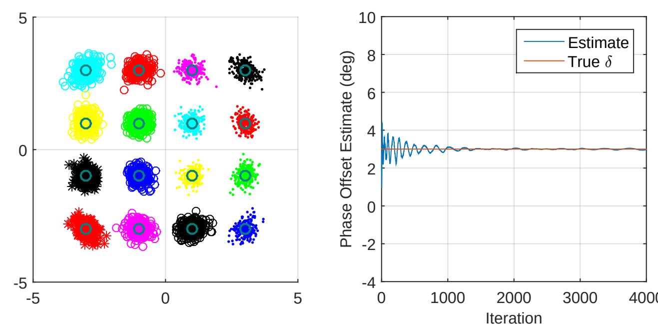

Fig. 5. Resulting compensated constellation (left) and frequency offset estimation performance (right) for sequential Bayesian estimator (9)-(13).

6. CONCLUSION

We have proposed an online Bayesian framework for blind estimation of the frequency offset and the parameters of an arbitrary signal constellation. Our approach leverages novel conjugate prior distribution theory for the von Mises and Gaussian distribution, which allows us to derive closed-form updates of the hyperparameters of the posterior distribution of the unknown parameters given streaming data. Simulations show that our estimation algorithm results in fast convergence and learning of the frequency offset, and significantly outperforms heuristic online frequency offset estimators.

7. REFERENCES

[1] U. Mengali and A. N. D'Andrea, *Synchronization techniques for digital receivers*, Applications of Communications Theory. Plenum Press, New York, London, 1997.

[2] G. Vazquez and J. Riba, "Non data-aided digital synchronization," in *Trends in Single and Multi-User Systems (Vol. II)*, G. B. Giannakis and Y. Hua, Eds. Prentice-Hall, Upper Saddle River, NJ, 2000.

[3] H. Fu and P. Y. Kam, "MAP/ML Estimation of the Frequency and Phase of a Single Sinusoid in Noise," *IEEE Transactions on Signal Processing*, vol. 55, no. 3, pp. 834–845, March 2007.

[4] T. Ghirmai, M. F. Bugallo, J. Miguez, and P. M. Djuric, "A Sequential Monte Carlo Method for Adaptive Blind Timing Estimation and Data Detection," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2855–2865, August 2005.

[5] T. Tsiligkaridis and K. W. Forsythe, "Adaptive Low-Complexity Sequential Inference for Dirichlet Process Mixture Models," *Preprint, arXiv: 1409.8185*, September 2014.